Benchmark Assessment Models

NOAA FISHERIES
NW Fisheries Science Center

Ian Taylor (NWFSC)

Overview

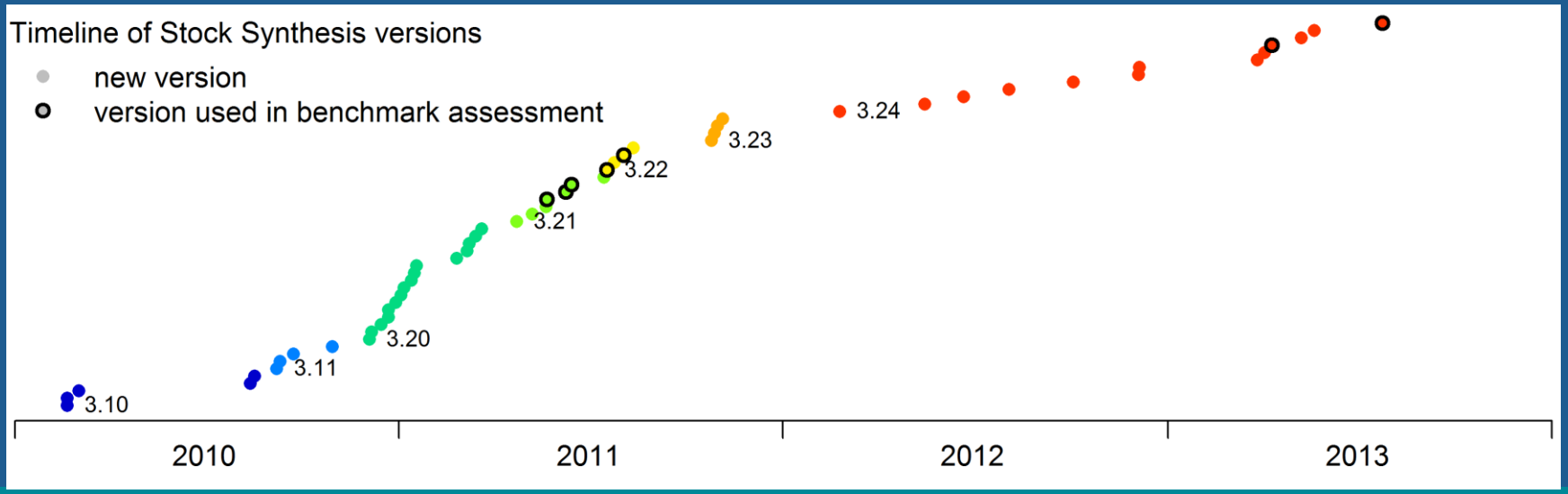
- Stock Synthesis
- Data types used in benchmark assessments
  - Indices
  - Age and length compositions
  - Discards
- Modeling growth
- Complexity of benchmark assessments
- Exploring alternative models

Stock Synthesis (SS)

- The modeling tool for all west coast groundfish benchmark assessments
  - Not required, but makes meeting Terms of Reference easier
- Very flexible and widely used (60+ stocks worldwide)
- New features added frequently
  - 72 versions since December 2009
  - 7 used for benchmark assessments
- NWFSC staff very familiar with software
  - Rick Methot has been responsive to requests for changes
  - Assessment team provides ideas and code for new features
  - Each new SS version tested by assessment team against a set of test models

Stock Synthesis (SS)

- The modeling tool for all west coast groundfish benchmark assessments
  - Not required, but makes meeting Terms of Reference easier
- Very flexible and widely used (60+ stocks worldwide)
- New features added frequently
  - 72 versions since December 2009
  - 7 used for benchmark assessments



SS as Common Platform

- Makes collaboration efficient
- Makes reviewing more efficient
- Fosters development of shared R code
- Potential drawback: assessment authors don't add new model features themselves
  - Could be a good thing

Data types

- All benchmark assessments are fit to at least 1 index of abundance
- All benchmark assessments have length and/or age data

Indices of abundance

- Assumed to have log-normal error distribution
- Catchability never fixed
- Selectivity typically estimated
- Additional variance parameter estimated when necessary

Indices of abundance

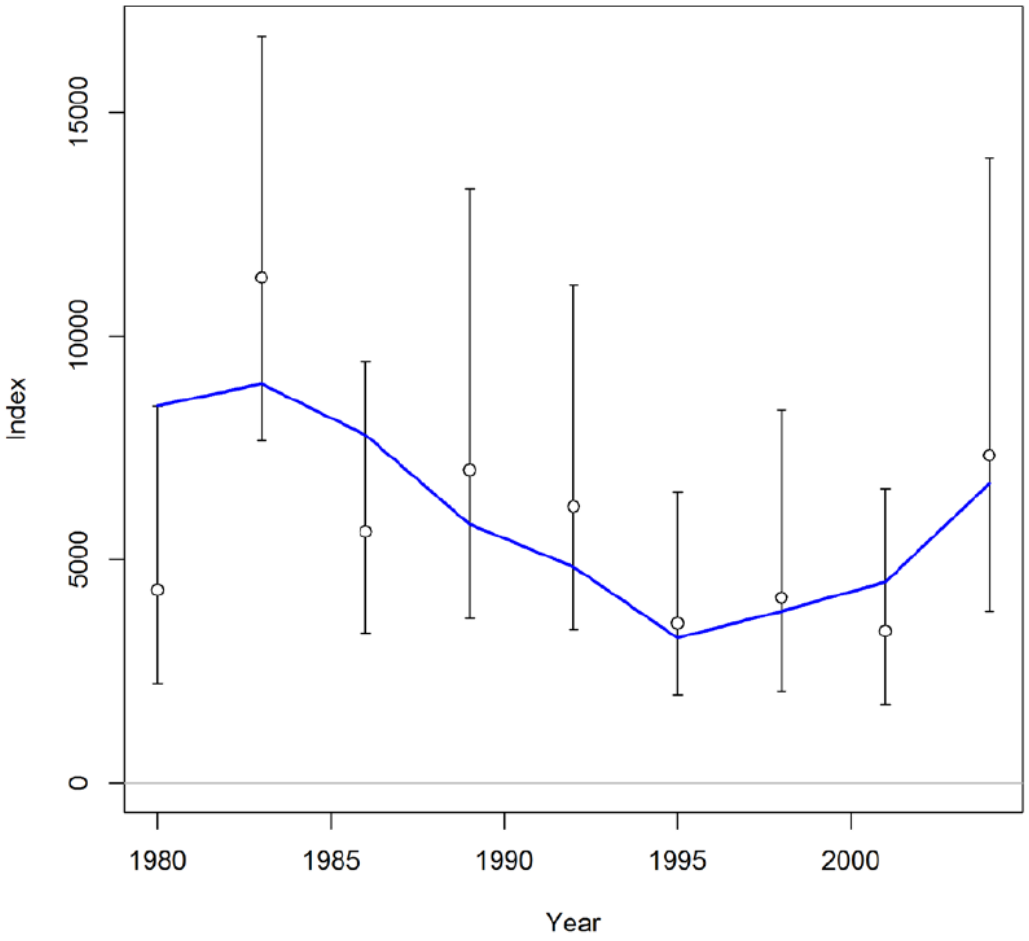


Figure 49: Observed and expected values of darkblotched rockfish biomass index (mt) for the AFSC shelf survey.

Figures from Darkblotched assessment

Length compositions

- Length data typically available for most fleets in benchmark assessments
- Availability of length data is impediment to doing more benchmark assessments

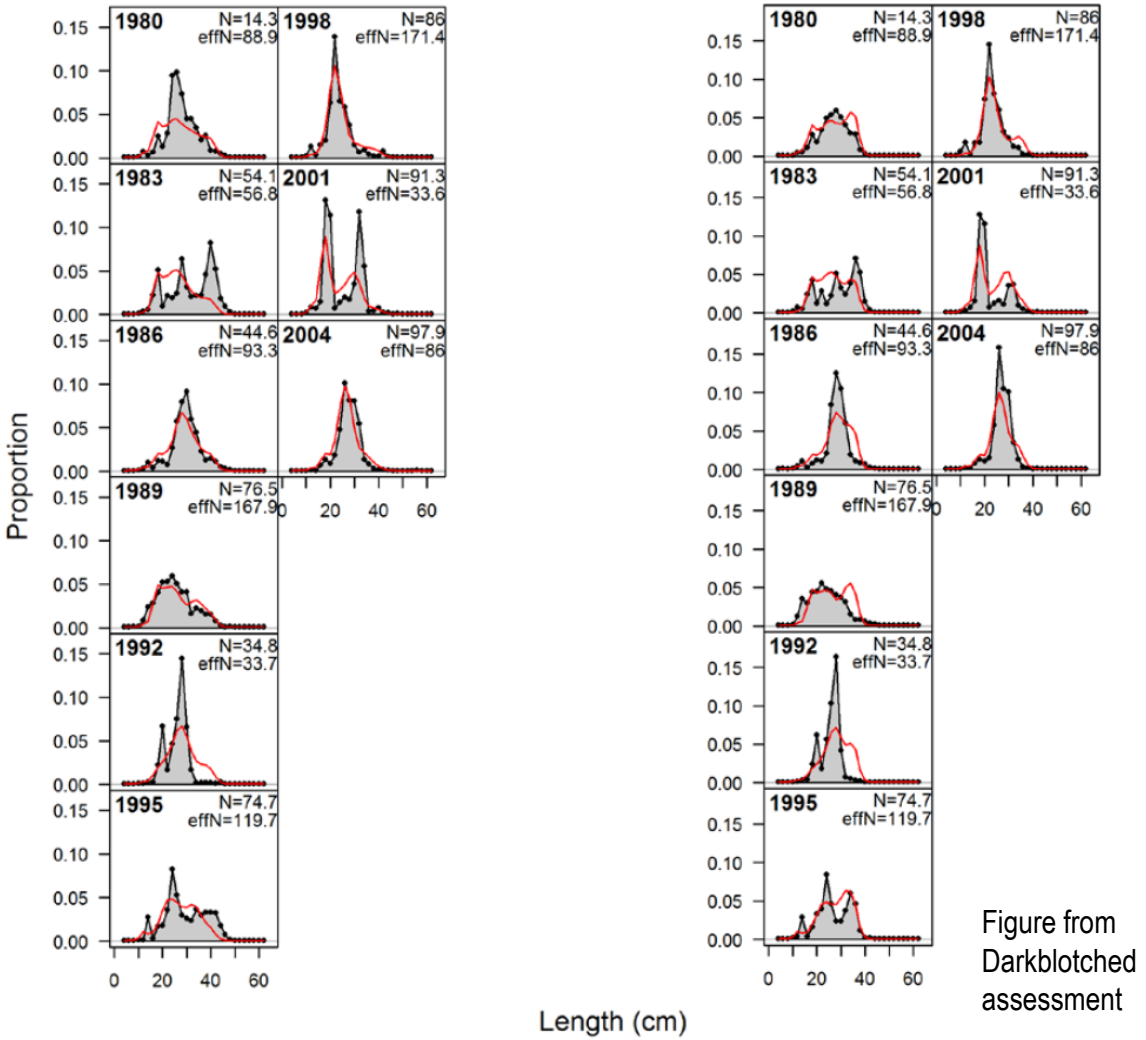


Figure from Darkblotched assessment

Figure 62: Fit to length-frequency distributions of female (left panel) and male (right panel) darkblotched rockfish from the AFSC shelf survey, by year.

NOAA FISHERIES

Age data

- Production ageing not available for all species
  - Assessments without age data typically have externally estimated growth curve
  - Benchmark assessments without informative age data usually assigned to Category 2
- Age data typically represented as conditioned on length bin (“conditional length-at-age data”)

Conditional age-at-length data

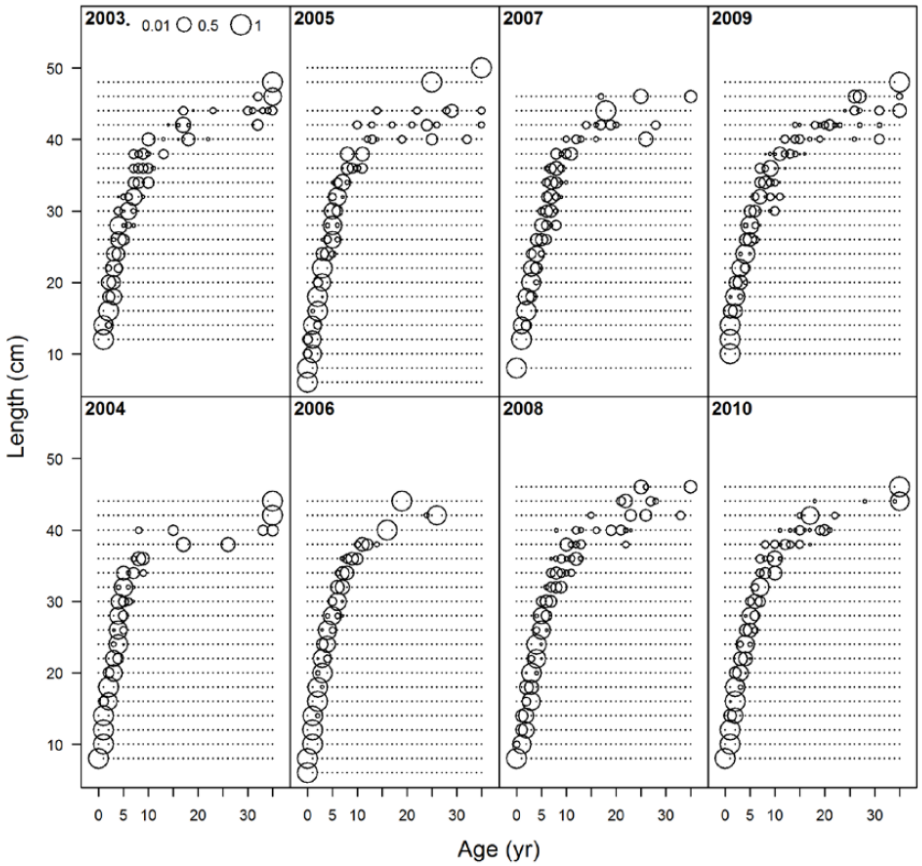


Figure 28: Conditional age-frequency distributions for female darkblotched rockfish from the NWFSC shelf-slope survey.

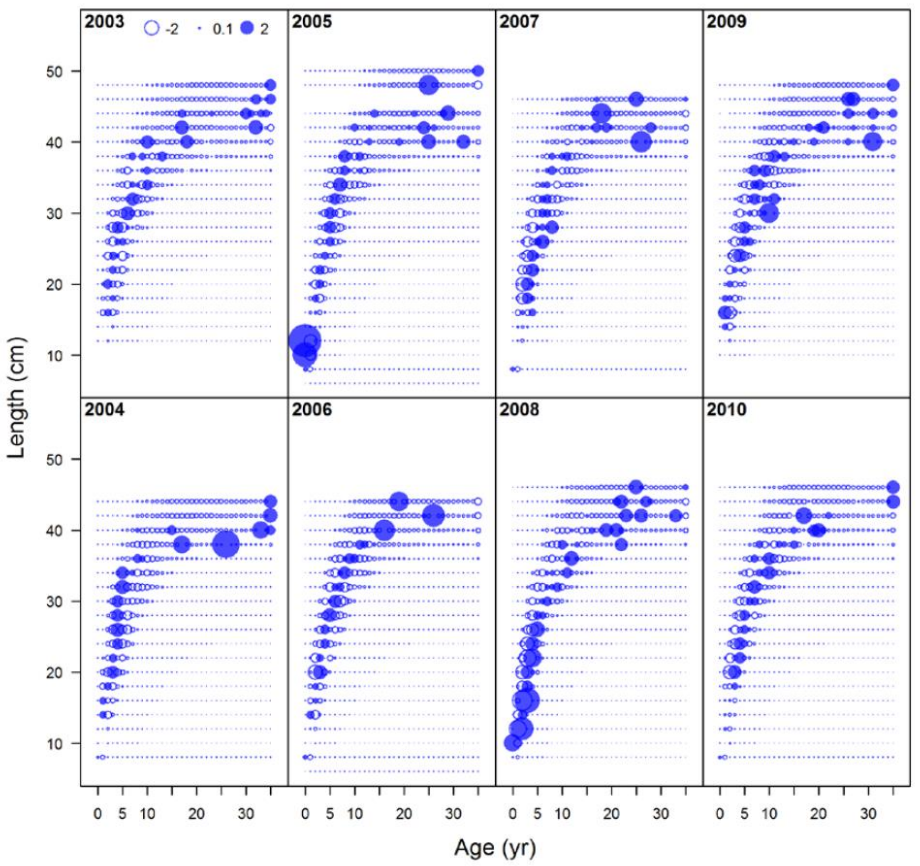


Figure 98: Pearson residuals for the fit to conditional ages-at-length compositions of female darkblotched rockfish from the NWFSC shelf-slope survey.

Figures from Darkblotched assessment

Conditional age-at-length data

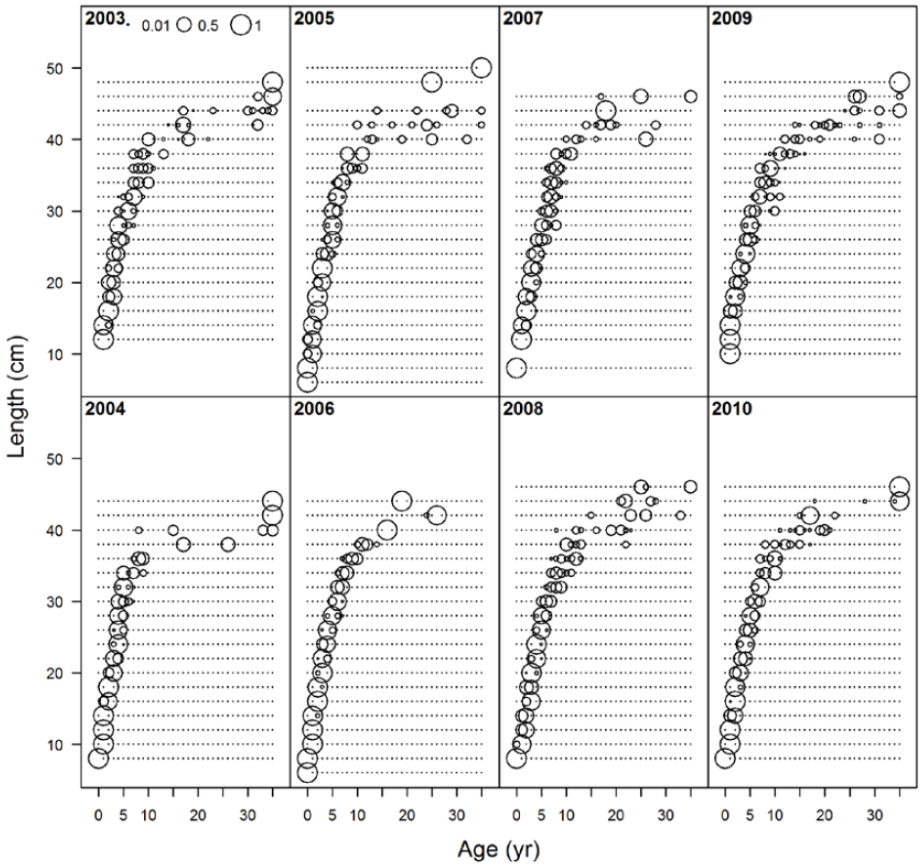


Figure 28: Conditional age-frequency distributions for female darkblotched rockfish from the NWFSC shelf-slope survey.

- Avoids double counting samples (orthogonal)
- Avoids influence of non-random subsampling for ages
- Avoids influence of ageing error on observed length at age
- Facilitates estimation of growth curve and variability around growth curve

Figures from Darkblotched assessment

Conditional age-at-length data

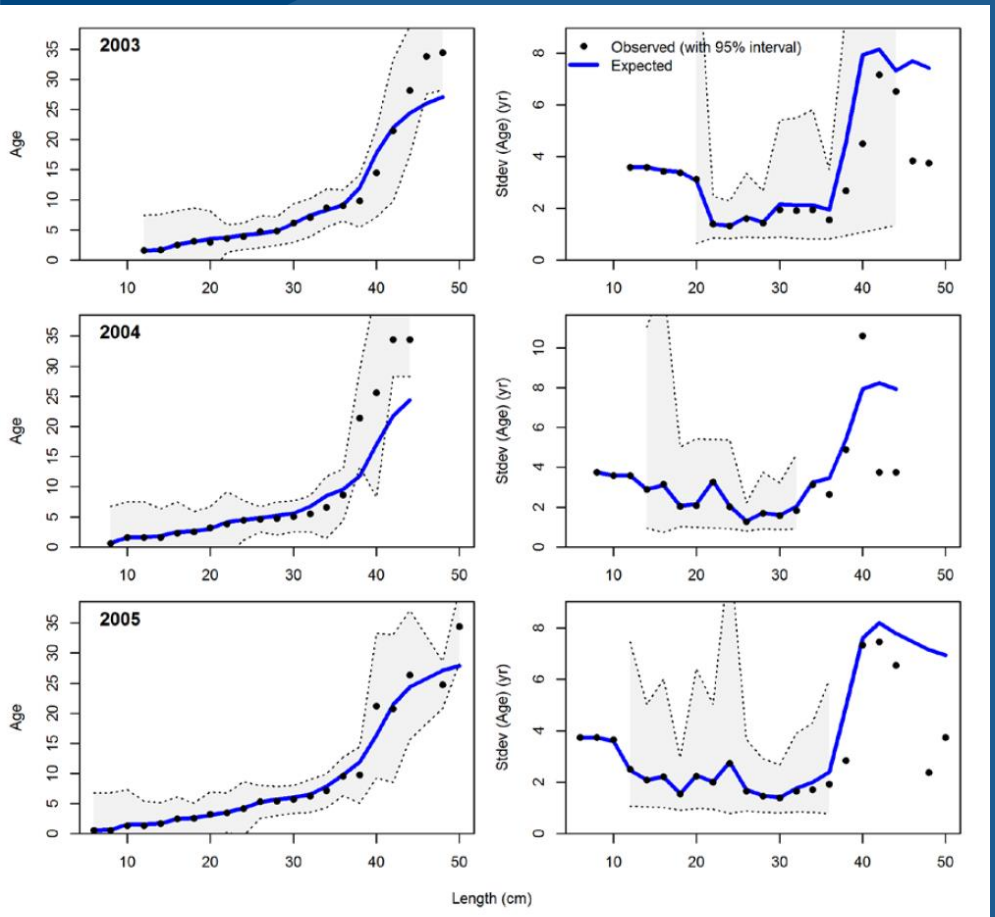


Figure 96: Fit to conditional ages-at-length compositions of female darkblotched rockfish from the NWFSC shelf-slope survey.

- Avoids double counting samples (orthogonal)
- Avoids influence of non-random subsampling for ages
- Avoids influence of ageing error on length-at-age
- Facilitates estimation of growth curve and variability around growth curve
- Challenge: interpretation

Figures from Darkblotched assessment

Marginal age data

- Not included in likelihood when conditional age-at-length data are used
- Useful for identification of patterns
- Sometimes easier to interpret

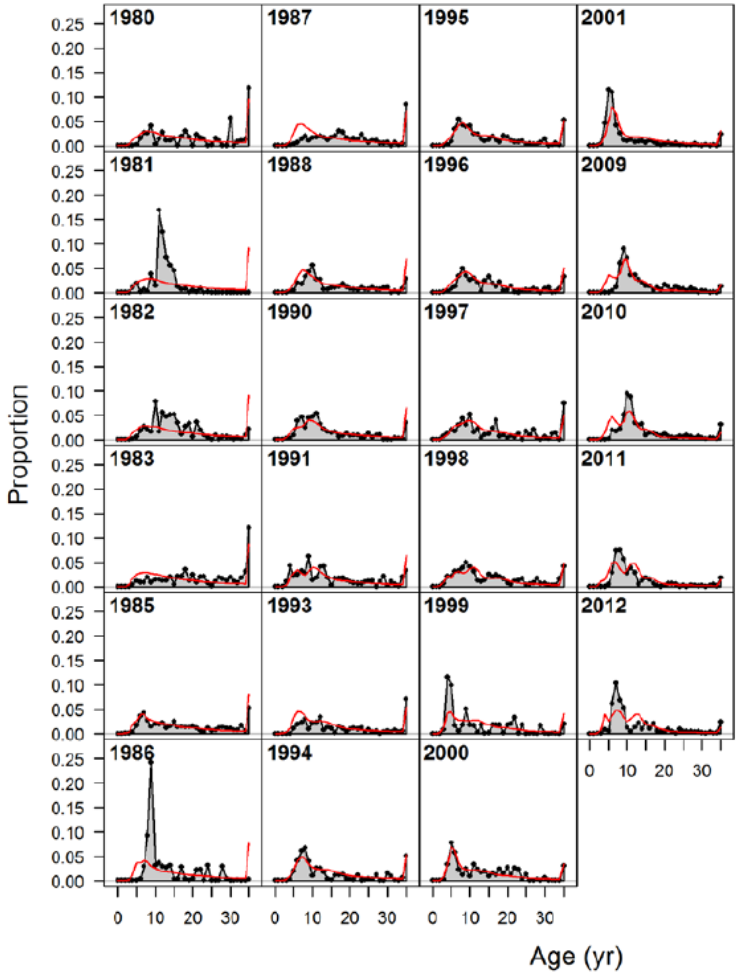


Figure from Darkblotched assessment

Figure 81: Implied fit to "ghost" marginal age compositions for female darkblotched rockfish from the domestic trawl landings. Age data from these years were not explicitly used in the assessment. Fits are provided for evaluation only, but not included in the model likelihood.

Age data for hake

- Pacific Hake population dynamics driven by very large recruitment events
- Large samples sizes of ages, extended back many years
- Highly variable growth difficult to model parametrically
- Current assessment uses empirical weight-at-age data but no length compositions

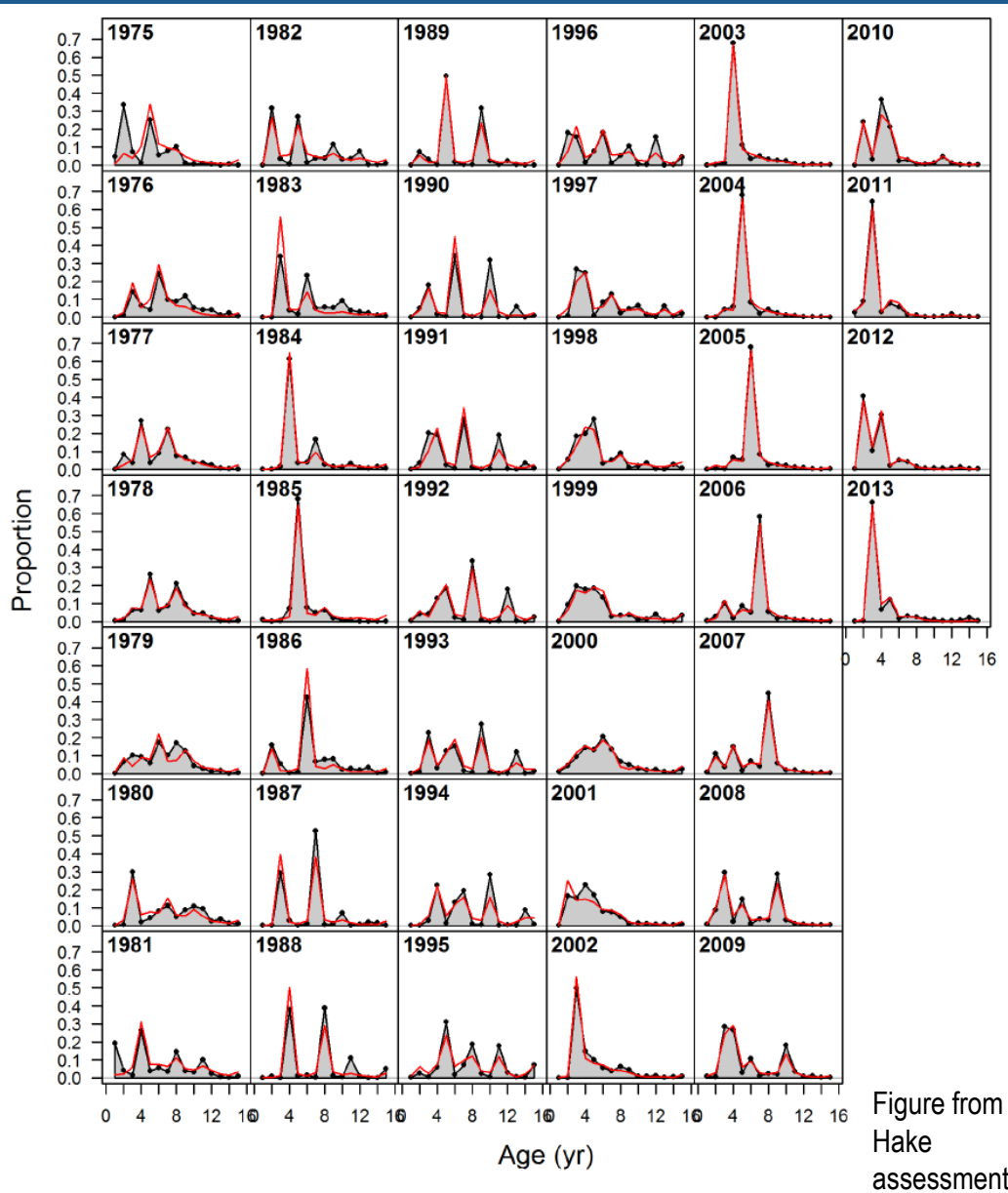


Figure 26: Base model fit to the observed fishery age compositions.

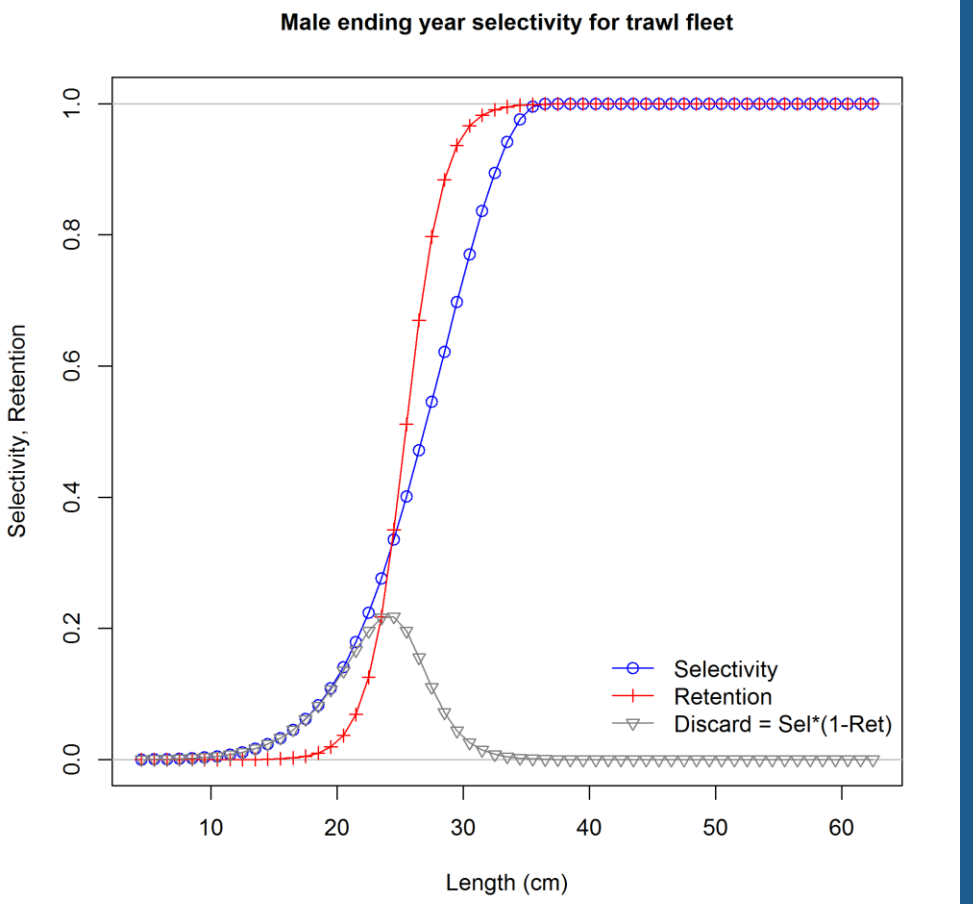
Figure from Hake assessment

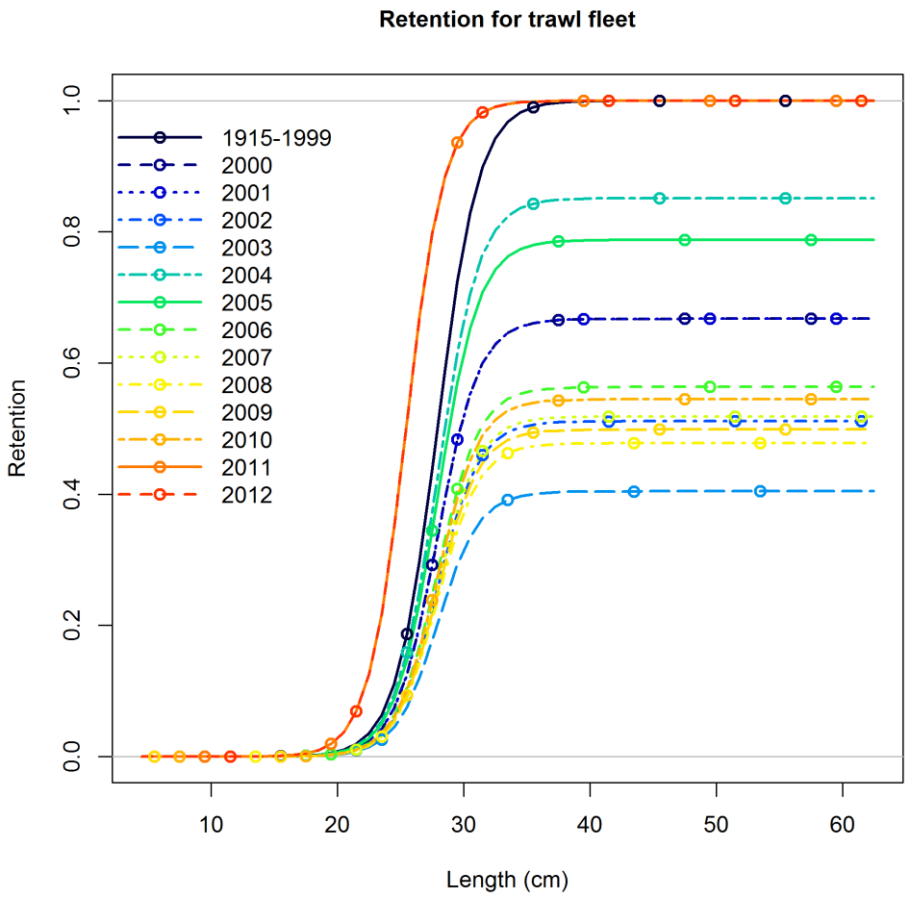
NOAA FISHERIES

Modeling Discards

- Discards available from West Coast Groundfish Observer Program from 2003 onward
  - Smaller sampling programs occurred in earlier years
- Patterns of show discard rates changing frequently
  - Changes in management have big impact on discards
- Discards typically modeled through a length-based retention function applied on top of selectivity
  - Accounts for changes in length composition of the population
  - Flexible for modeling changes over time and periods with no data

Modeling Discards





Figures from Darkblotched model

Modeling Discards

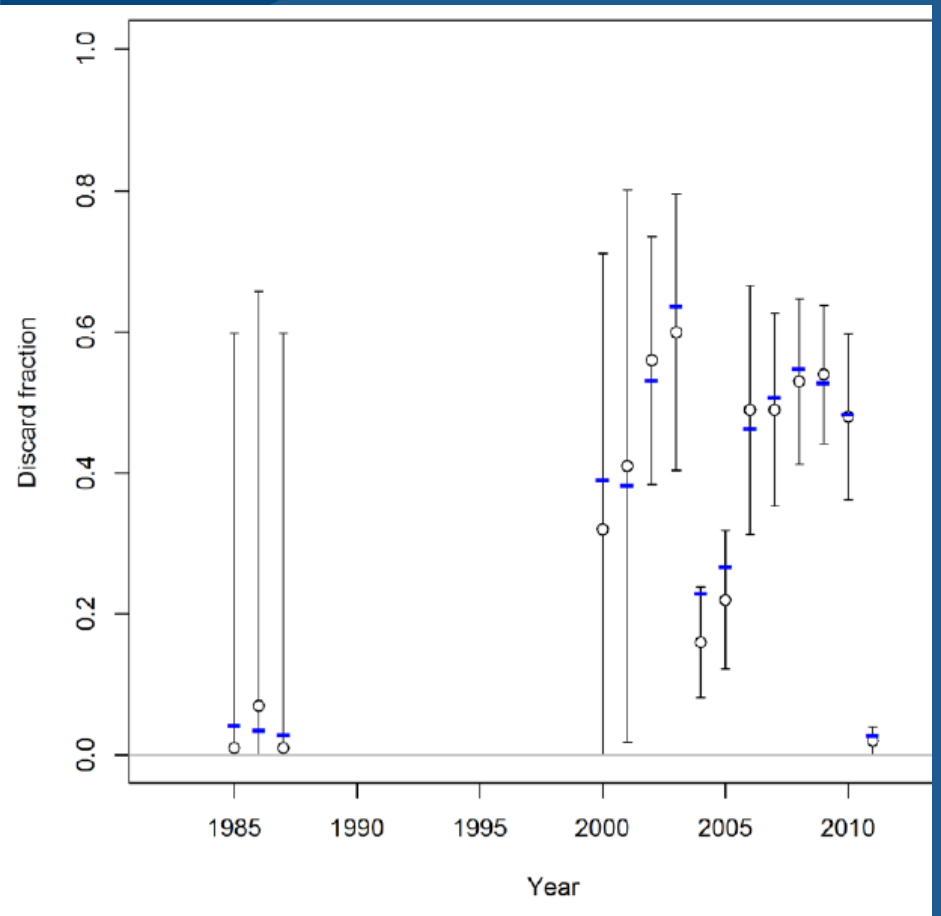


Figure 112: Fit to the discard ratio data of the domestic trawl fishery.

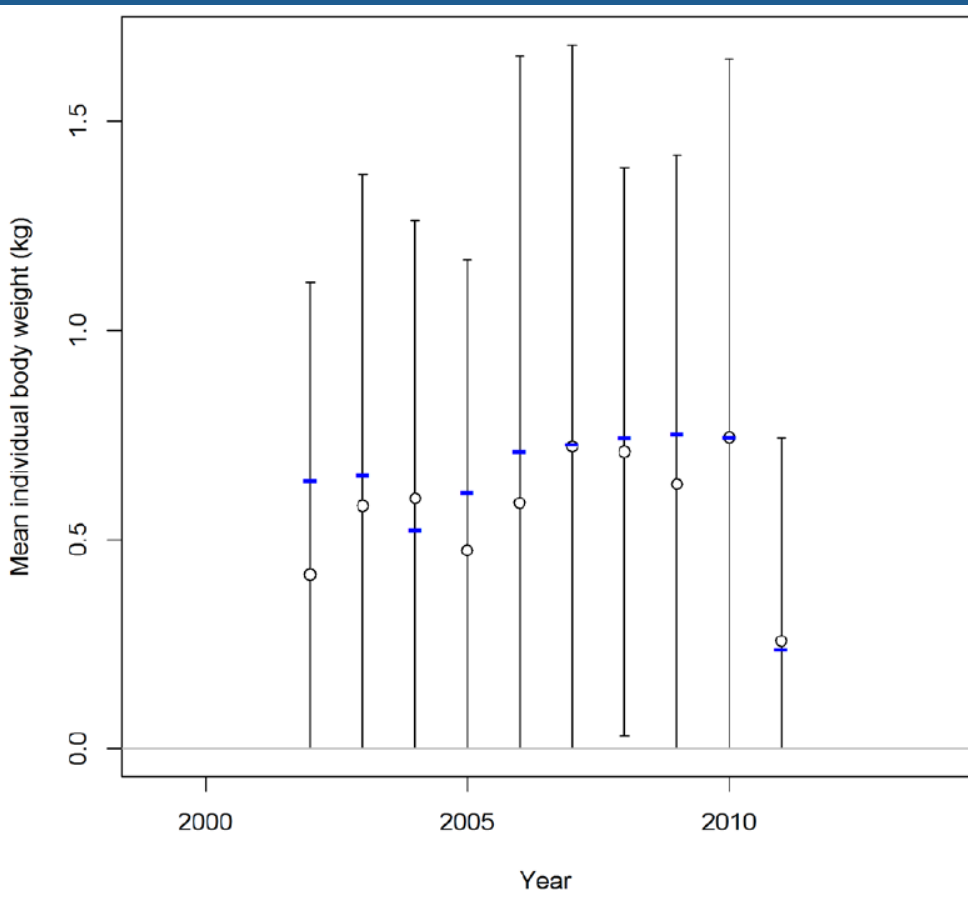


Figure 115: Fit to the mean body weight data for the domestic trawl fishery discard.

Figures from Darkblotched assessment

Additional Data Types Available

- Environmental variables can be linked to any parameter in SS
  - Recruitment, growth, mortality
  - Option not commonly used in production assessments (see Agenda item F.2)
- Mean body weight
  - Often used for discard data
- Mean length or weight at age
  - Problematic in the presence of ageing error
- Tagging data
  - Not available for most species
  - Not generally used in production assessments

Modeling growth

- Growth typically estimated within the assessment model for benchmark assessments
- Can accounts for influence of length-based selectivity on observed lengths
- Allows uncertainty in growth rates to be included in results

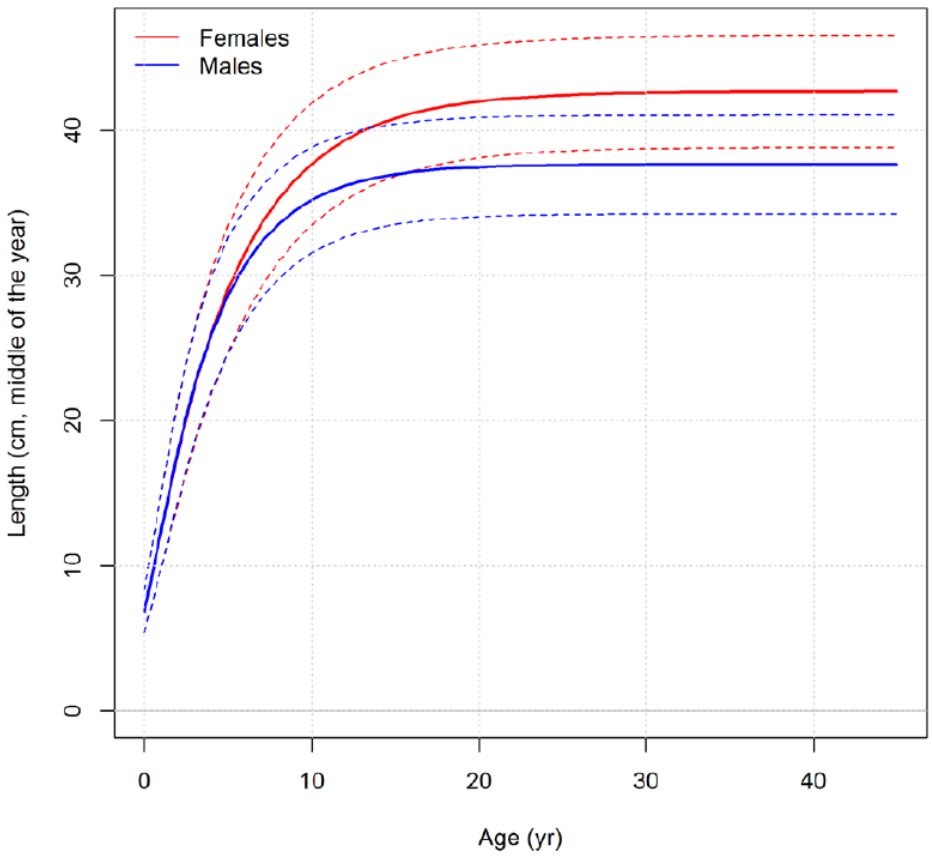


Figure 44: Growth curves for females and males of darkblotched rockfish used in the assessment model.

Figure from Darkblotched assessment

Modeling growth

- Hake assessment is only one to use empirical weight-at-age data
- Requires more age samples
- Only works with age-based selectivity
- Choice driven by difficulty in modeling variability among years and fast growth within each year

Figure from Hake assessment

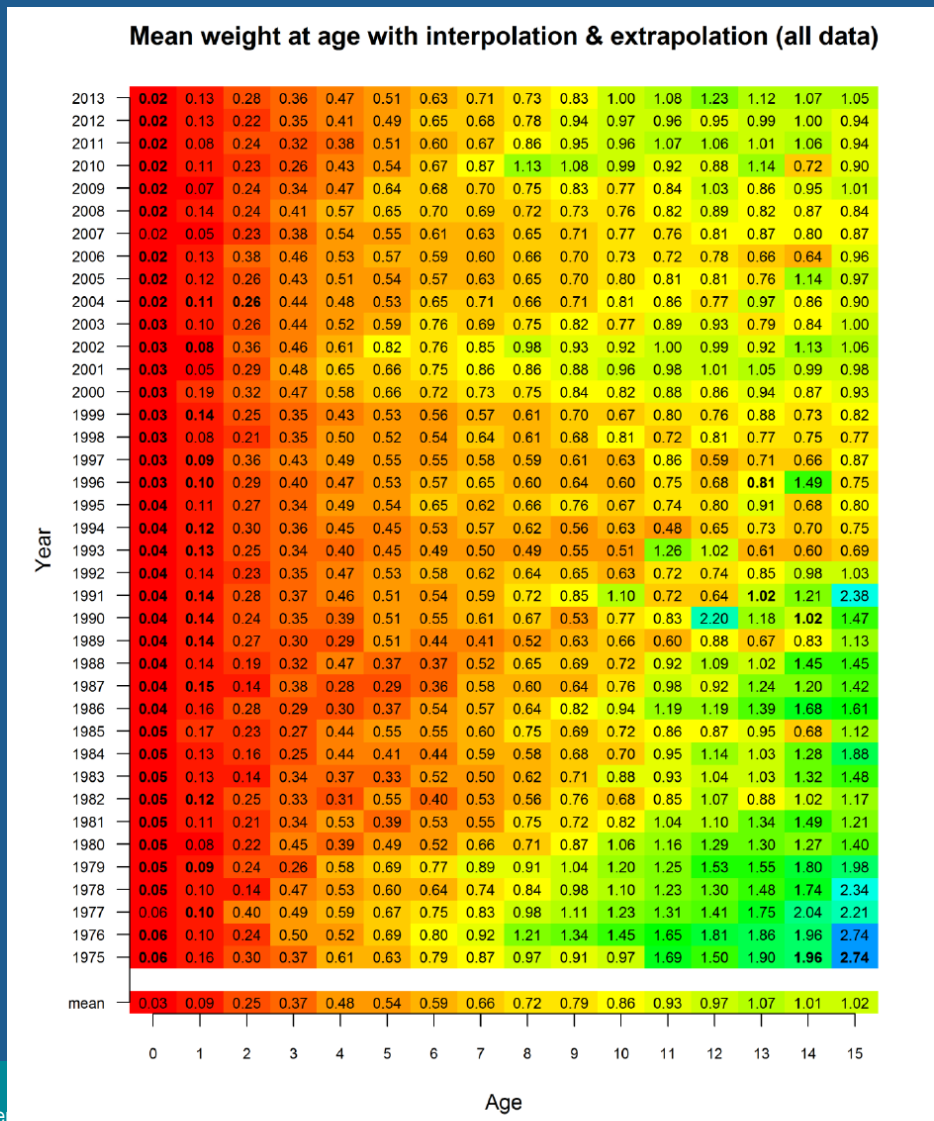
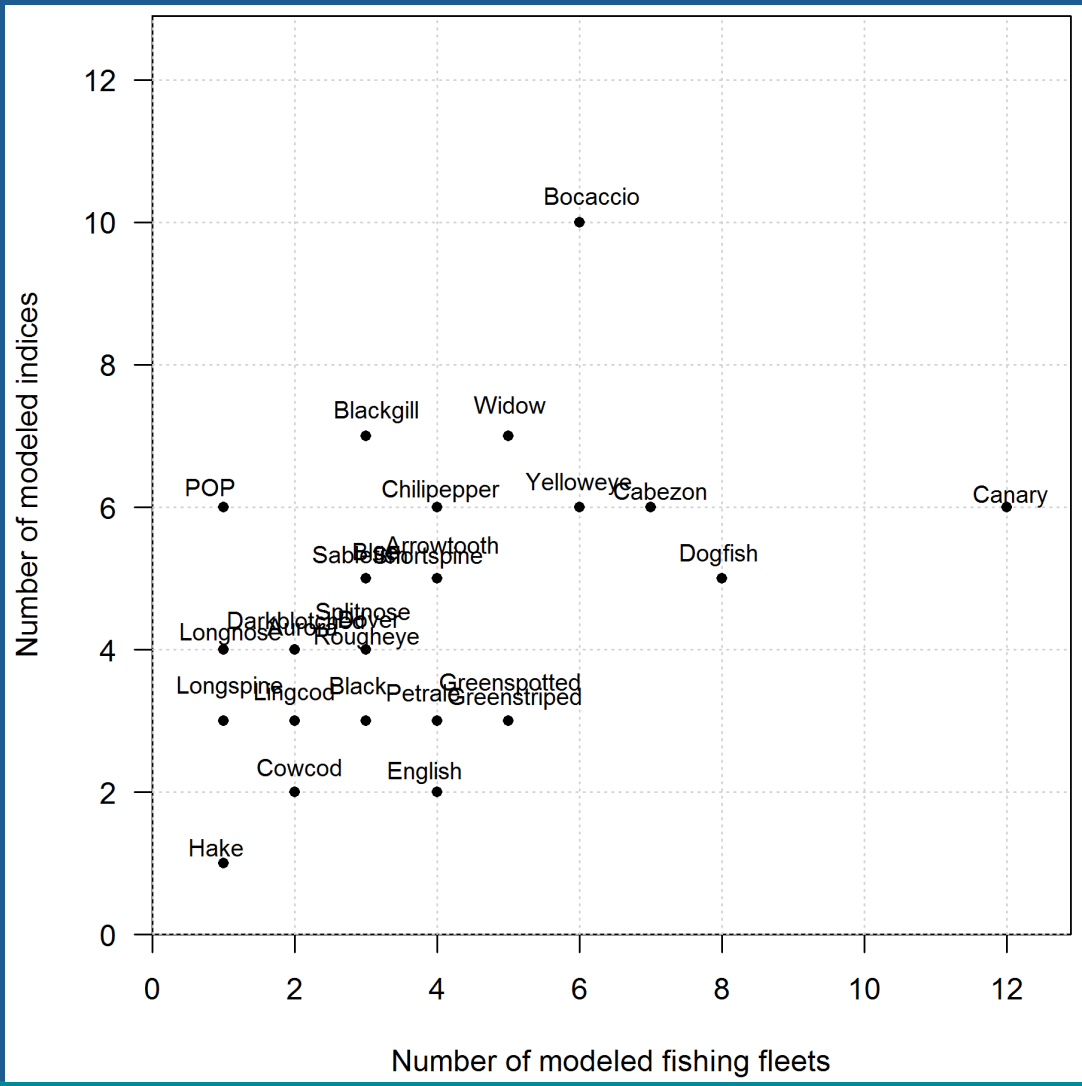


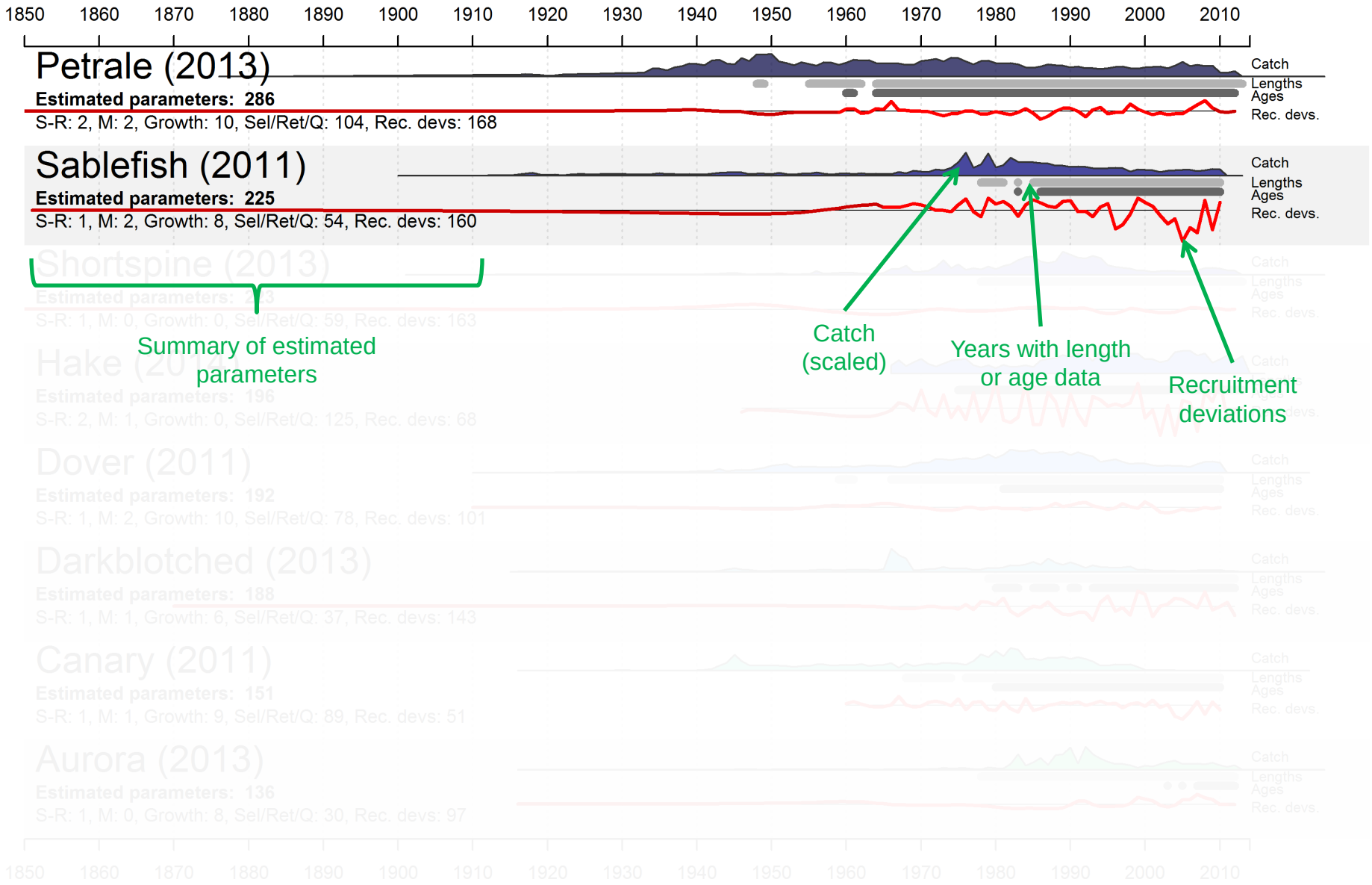
Figure 17: Empirical weight-at-age (kg) used in the assessment. Numbers shown in bold were interpolated or extrapolated from adjacent years.

Fleet structure

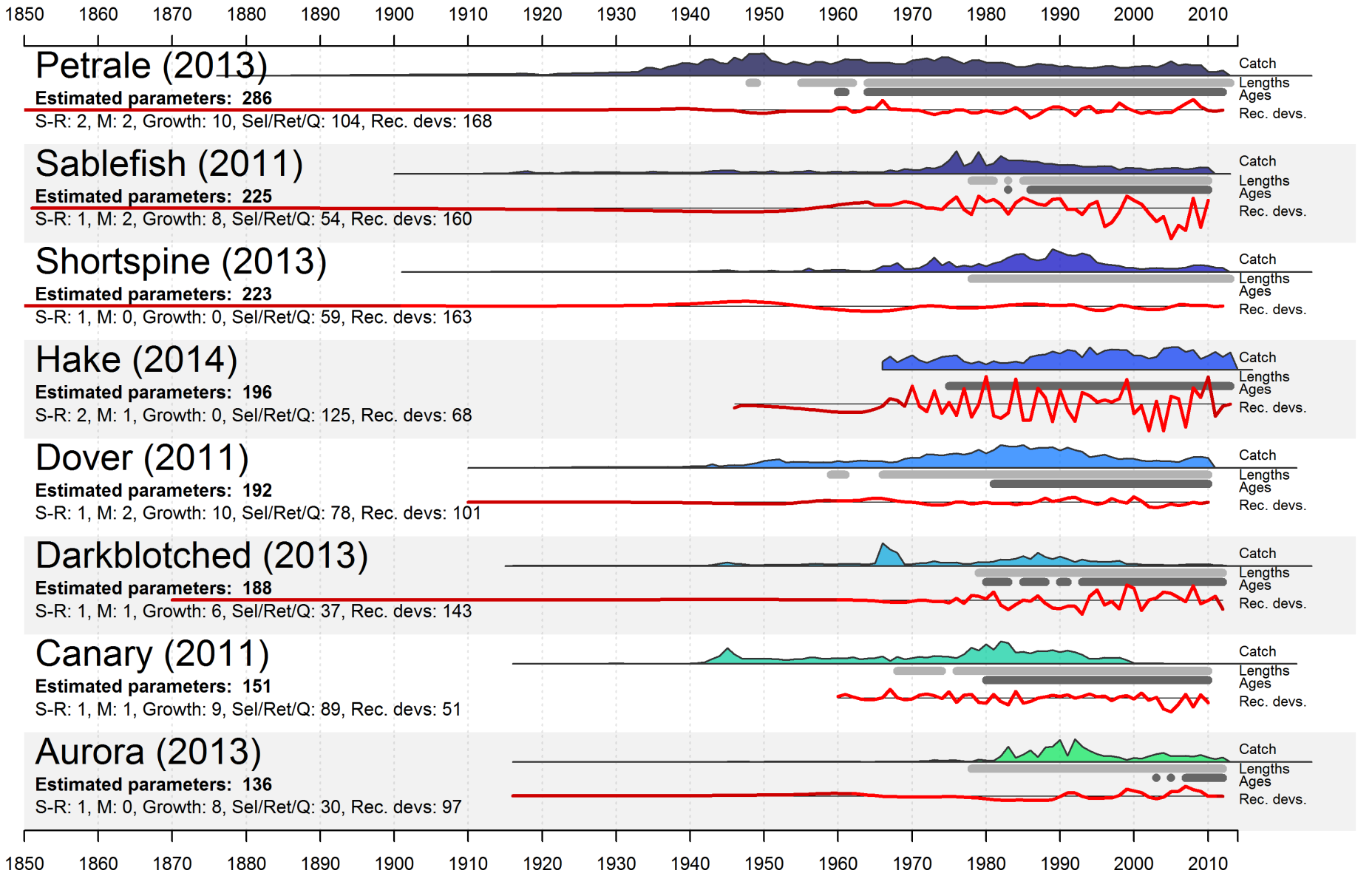
- Fleets are typically chosen based on combinations of gear, area, or time-period



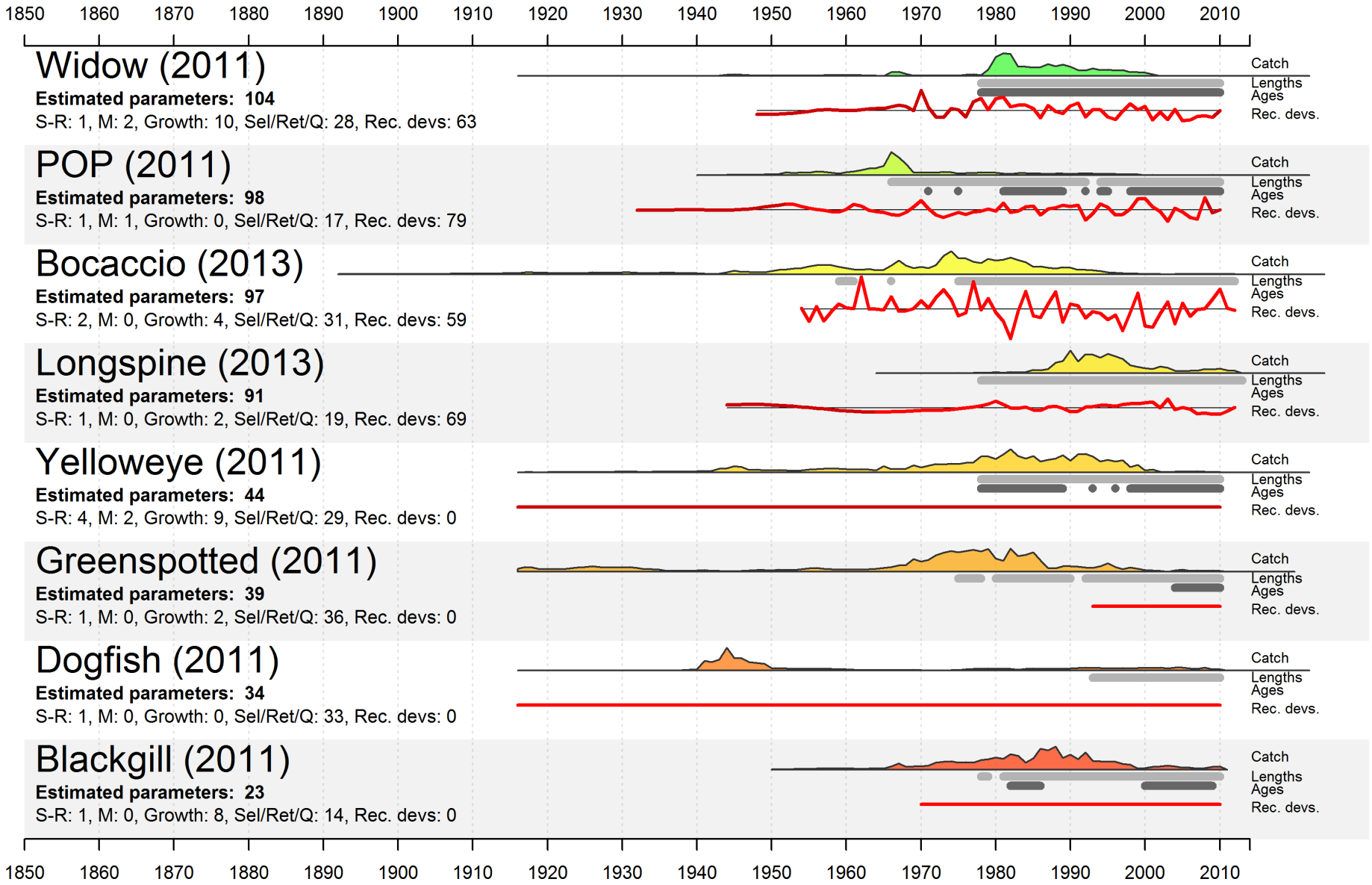
West Coast Benchmark Assessments 2011-2014 (pg. 1/2)



West Coast Benchmark Assessments 2011-2014 (pg. 1/2)



West Coast Benchmark Assessments 2011-2014 (pg. 2/2)



Parameters Estimated

- Equilibrium recruitment (R_0, $N = 1$)
- Other stock-recruit parameters ($N = 0–3$)
- Annual recruitment deviations ($N = 0–168$)
- Selectivity, retention, catchability ($N = 14–125$)
- Growth ($N = 0–10$)
- Natural mortality ($N = 0–2$)
- Initial equilibrium fishing mortality ($N = 0–1$)

Stock-recruit relationships

- Steepness parameter most influential for models with no estimated recruit. devs.
- Practices for modeling recruitment continue to evolve
- Topic of much ongoing research

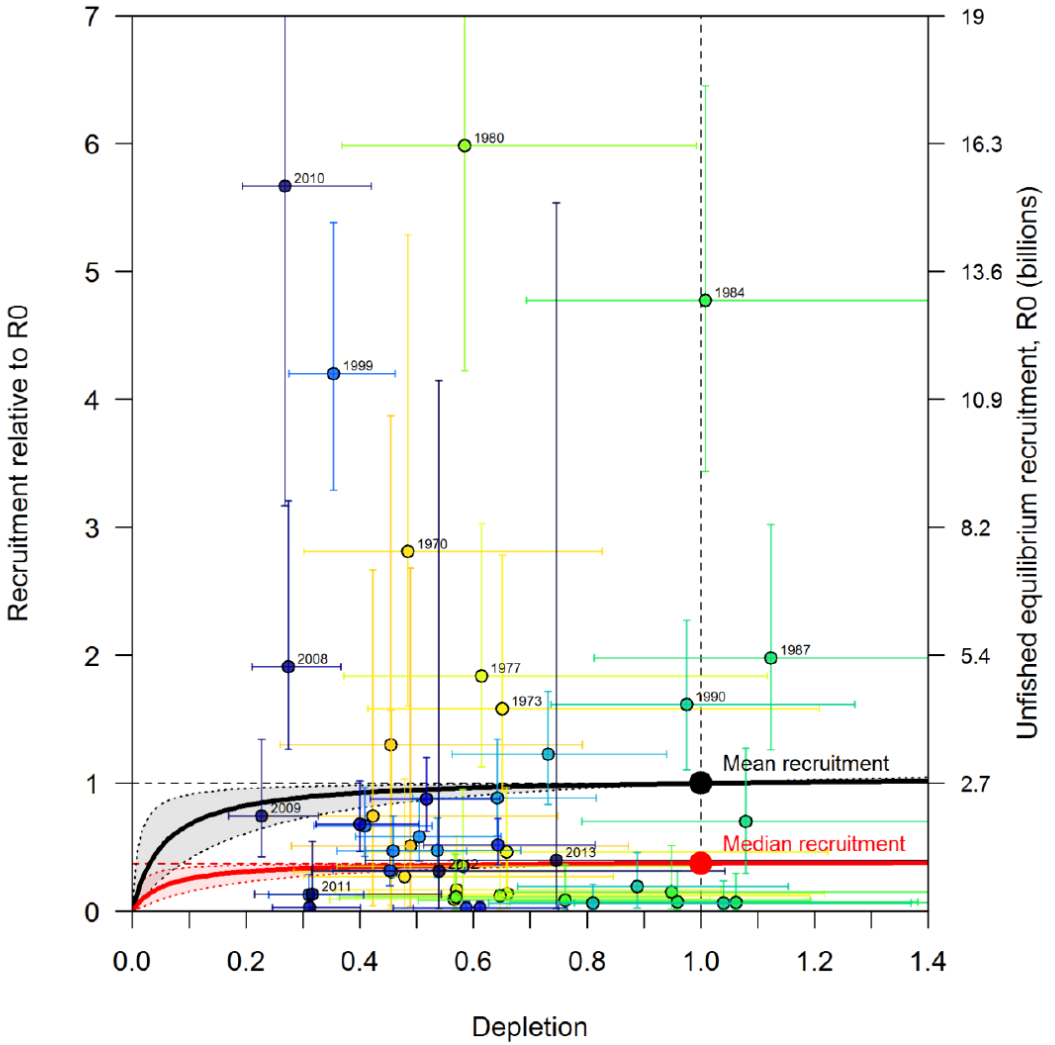


Figure 36: Estimated stock-recruit relationship for the base model with median predicted recruitments and 95% posterior credibility intervals. The thick solid black line indicates the central tendency (mean) and the red line the central tendency after bias correcting for the log-normal distribution (median).

Figure from Hake assessment

MLE vs. MCMC

- Hake assessment is only benchmark where MCMC is fast enough
- Comarisons discussed in Stewart et al. (2013).

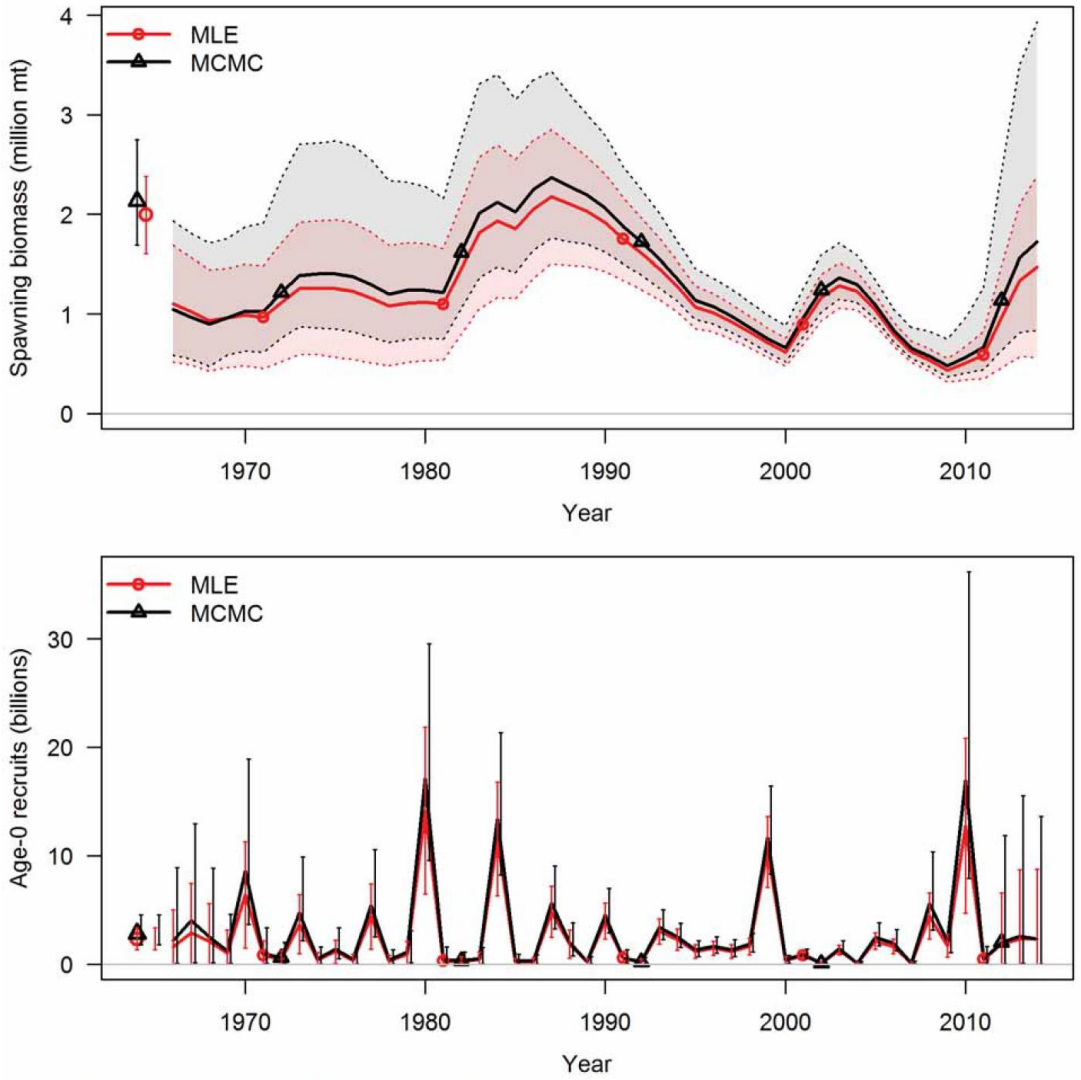


Figure 41: A comparison of MLE estimates with 95% confidence intervals determined from asymptotic variance estimates (red) to the median of the posterior distribution with 95% credibility intervals (black).

Figure from Hake assessment

Exploring alternative models

- Sensitivities, retrospective analyses, and likelihood profiles typically conducted in benchmark assessments
- Help reveal influence of alternative model structures, data types, and assumptions made in base model
- Help focus discussion on choices that are most influential

Sensitivity analysis

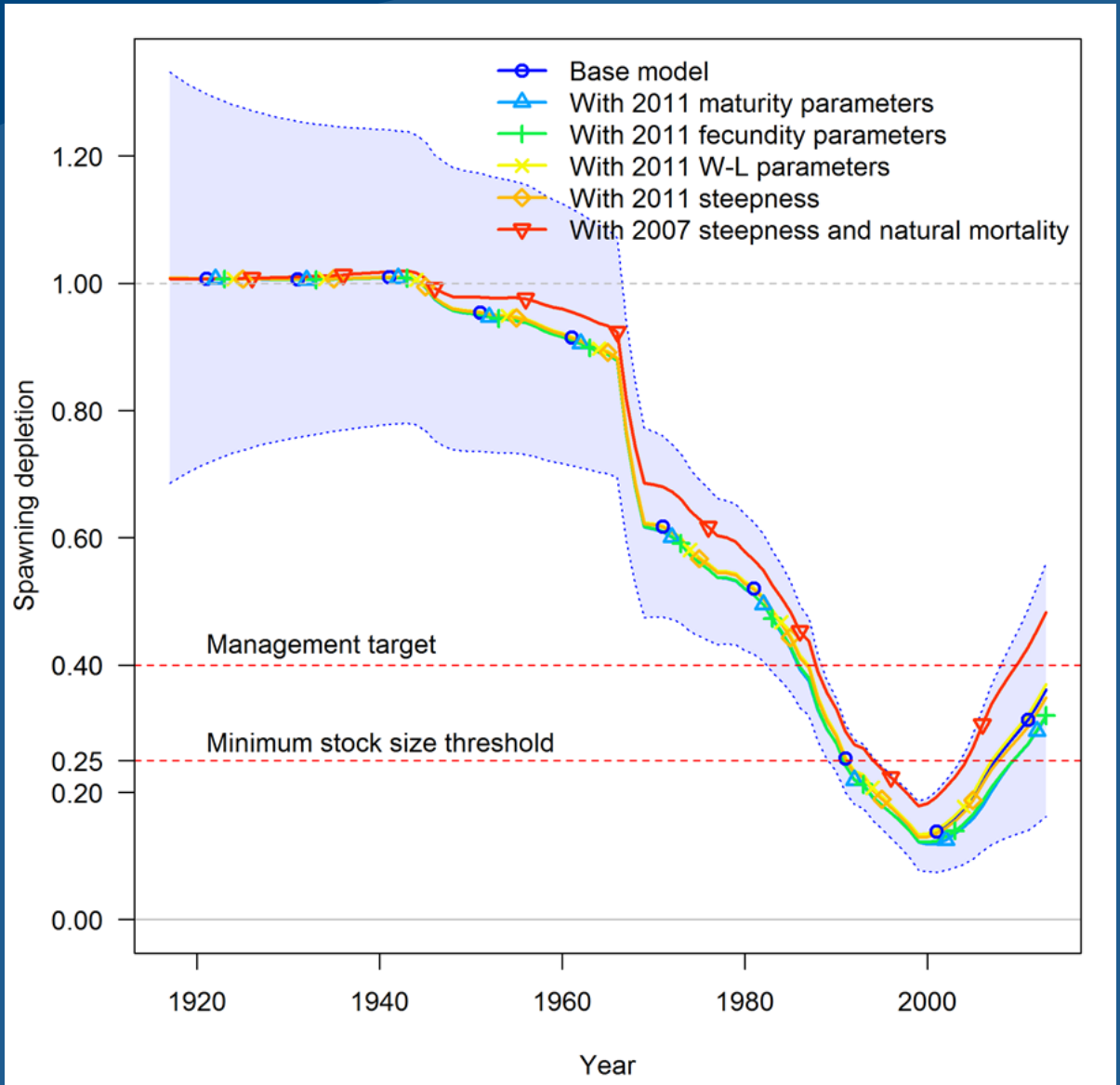


Figure from Darkblotched assessment

Retrospective analysis

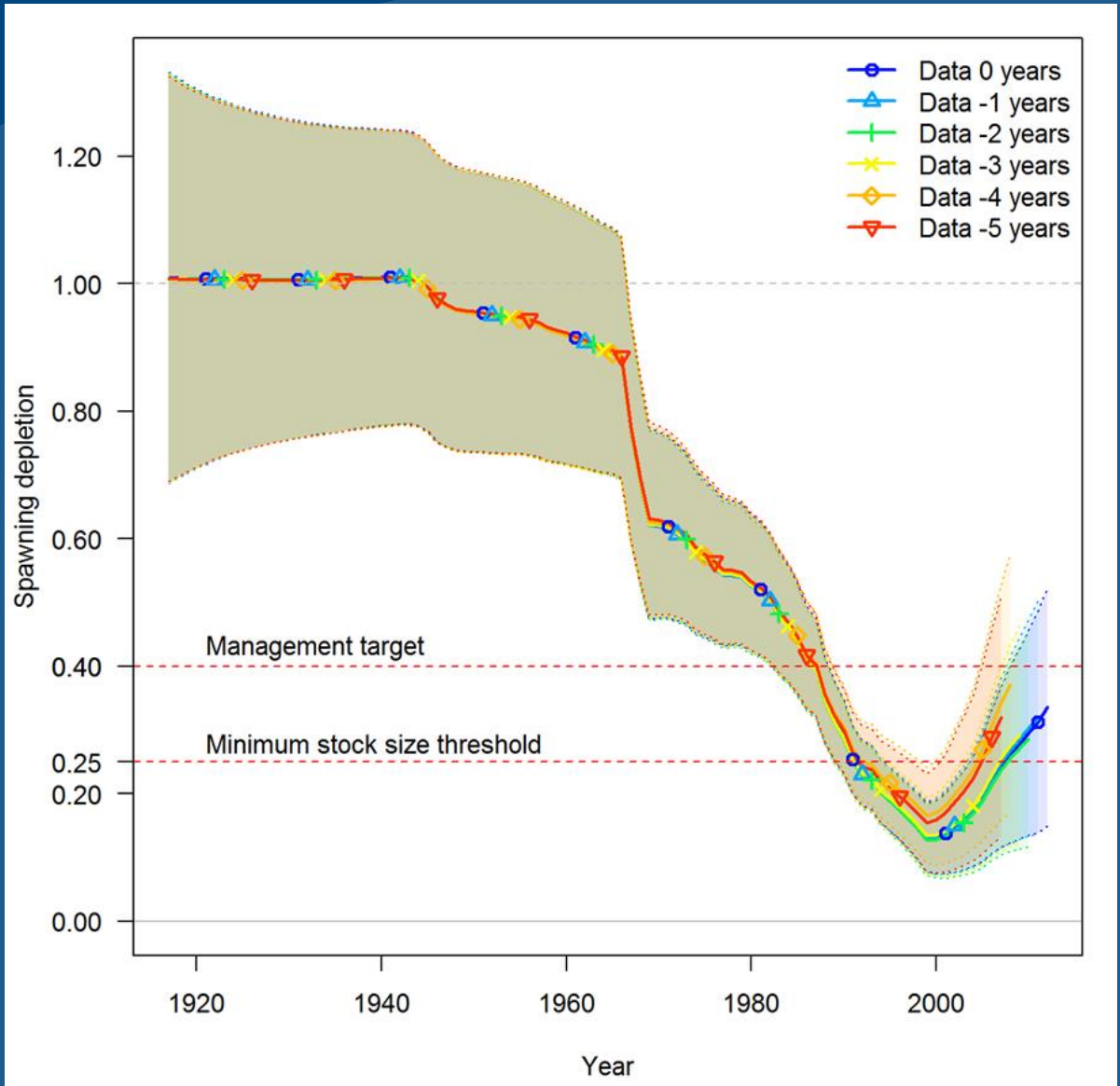


Figure from Darkblotched assessment

Likelihood profile analysis

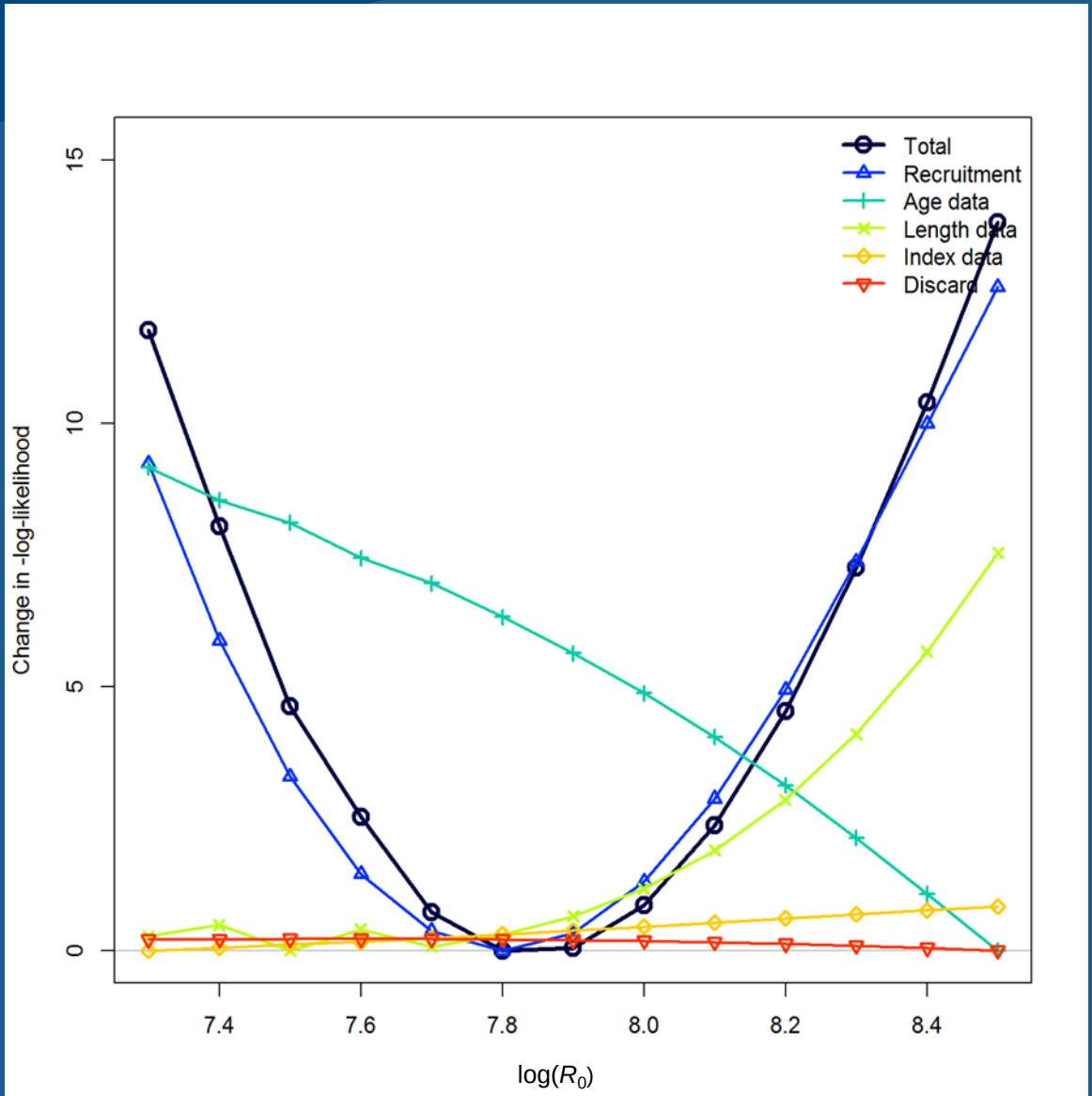


Figure from Darkblotched assessment

Decision table

			State of nature					
			Low *Female M=0.036*		Base case Female M=0.05		High *Female M=0.082*	
Management decision	Year	Catch (mt)	Spawning output (million eggs)	Depletion	Spawning output (million eggs)	Depletion	Spawning output (million eggs)	Depletion
Catch calculated using SPR of 71.9% applied to the base model	2013	223	607	18%	1,214	36%	3,606	82%
	2014	240	648	19%	1,294	39%	3,770	85%
	2015	252	688	20%	1,374	41%	3,922	89%
	2016	260	722	21%	1,441	43%	4,032	91%
	2017	266	751	22%	1,496	45%	4,101	93%
	2018	271	776	23%	1,541	46%	4,135	94%
	2019	276	798	23%	1,578	47%	4,147	94%
	2020	280	821	24%	1,613	48%	4,150	94%
	2021	285	844	25%	1,646	49%	4,149	94%
	2022	289	867	25%	1,678	50%	4,146	94%
	2023	293	891	26%	1,709	51%	4,140	94%
	2024	297	915	27%	1,739	52%	4,133	94%

Table from Darkblotched assessment

Strengths

- Flexibility of Stock Synthesis works well for range of available data and needs of benchmark assessments
- Use of common platform has helped streamline the assessment process
- Assessment reports increasingly thorough, continue to meet expanding Terms of Reference

Challenges and Solutions

- Benchmark assessments are large, complex projects (and increasing in scope)
  - Larger assessment teams being used
  - Request limited number of benchmarks per cycle
- Data necessary for benchmark assessments unavailable for many stocks
  - Continue to pursue data-limited methods
- Ecosystem effects rarely included in assessments
  - See discussions under F.2
- Need to balance research into improving models vs. status-quo methods
  - Two-year assessment cycle should facilitate balance